e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:03/March-2024 Impact Factor- 7.868 www.irjmets.com

A REVIEW OF PREVIOUS RESEARCHES ABOUT MACHINE LEARNING THEORY IN PRODUCTION MANAGEMENT: FOCUSING ON REINFORCEMENT LEARNING

Choi Seong-Woo[*1]

[*1]Professor, Department Of Business Administration, Kyonggi University.

DOI: https://www.doi.org/10.56726/IRJMETS50593

ABSTRACT

Currently, production management of cutting-edge IT products takes place in a challenging environment of technological complexity and rapid market changes. In this environment, production control is increasingly important and has become a key factor in determining product quality, cost, and delivery time. High-tech IT product production lines are comprised of a complex of various processes and technologies, so problems that may arise at each stage can have a significant impact on the final product. Therefore, an efficient production management system plays an important role in minimizing risks that may occur during the manufacturing process, maintaining product quality, and reducing production costs. In this study, we review existing studies focusing on reinforcement learning theory among machine learning theories in the field of production management, especially scheduling, and suggest limitations of those studies and future research directions.

Keywords: Machine Learning, Reinforcement Learning, Production Management, Scheduling .

I. INTRODUCTION

Nowadays, the manufacturing of cutting-edge IT products takes place in a challenging environment that requires a high level of technical complexity and rapid market changes. In this environment, the importance of production management is increasing, and it has become a key factor in determining product quality, cost, and delivery time. The production line of cutting-edge IT products involves a complex combination of various processes and technologies, so problems that may arise at each stage can have a significant impact on the final product. Therefore, an efficient production management system plays an important role in minimizing risks that may arise during the manufacturing process, maintaining product quality, and reducing production costs.

Currently, global competition is intensifying in the cutting-edge IT product manufacturing industry, and consumer demands are diversifying and increasing. Accordingly, production management must go beyond simply internal efficiency to increase customer satisfaction and provide flexibility to quickly respond to changes in the market. Additionally, environmental aspects must be considered for sustainable production, which means minimizing the environmental impact that may occur during the production process and using resources efficiently.

In this context, production management of cutting-edge IT products is recognized as an essential field that must be approached from a strategic perspective beyond simply managing the manufacturing process. Efficiency in production management plays a critical role in reducing product time to market, maximizing cost efficiency, and ensuring the quality of the final product. To this end, high-tech IT product manufacturers are introducing advanced production management technologies and systems, including predictive analysis, automation, and real-time monitoring using the latest technologies such as big data, artificial intelligence, and the Internet of Things (IoT).

In conclusion, production management in the manufacturing line of cutting-edge IT products is a key factor in successfully launching products to the market, maintaining competitiveness, and achieving sustainable growth. This makes an important contribution to manufacturers' ability to respond nimbly to market changes, meet customer needs, and gain competitive advantage. Therefore, the role and importance of production management is expected to continue to increase in the future.

Meanwhile, the application of machine learning in production management can appear in various forms. For example, in complex decision-making processes such as production scheduling, inventory management, quality control, and supply chain optimization, reinforcement learning models can derive optimal decisions based on real-time data. This can help improve production line flexibility, reduce manufacturing costs, and increase customer satisfaction.

The powerful utility of machine learning comes primarily from the adaptability and learning capabilities it provides. Manufacturing processes and environments are constantly changing, making it difficult to achieve optimal performance with fixed rules or single decision models. machine learning provides the ability to dynamically adapt to these changes, allowing production systems to respond more nimbly and operate more efficiently. Additionally, reinforcement learning is useful for modeling complex problems and finding optimal solutions in a multidimensional decision space.

In actual applications, machine learning has already proven its effectiveness in production line optimization, energy consumption minimization, failure prediction, and maintenance scheduling. These examples provide clear evidence that machine learning has strong effectiveness in solving problems faced in the field of production management. Ultimately, the application of machine learning makes production systems more intelligent, automated, and efficient, providing a foundation for companies to remain competitive and operate successfully in the market.

In the main text, representative studies using machine learning theory in production management will be described. In the conclusion, the limitations of machine learning theory in the existing production management field will be presented and suggestions on the direction in which machine learning theory should be utilized in the production management field in the future will be made. In particular, we would like to clarify in advance that among machine learning theories, this review study focuses on reinforcement learning.

II. REVIEW OF EXISTING RESEARCH ON REINFORCEMENT LEARNING

Machine learning can be divided into three fields: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning [1] is a technique that trains an algorithm with data given the correct answer (labeling) and then derives the correct answer for the data when an X factor is given. Typical problems include classification and regression. Classification learning is divided into binary classification problems that determine O/X [2] and multi-classification problems that determine objects using images [3]. The problem of classifying dogs and cats is a representative binary classification problem, and a widely known multiple classification problem involves training an algorithm using human handwriting from numbers 1 to 9 and then judging the number using the handwriting of multiple people [4]. A regression problem [5] is an algorithm that infers continuous values, such as patterns or trends, from given features. Representative examples include defect rate prediction and time series analysis. Unsupervised learning refers to a technique that finds and clusters similar shapes or patterns within given features using data for which the correct answer (labeling) is not given. The most widely known unsupervised learning techniques include K-means Clustering [6] and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) [7].

Lastly, there is reinforcement learning [8], which is the topic of this paper. Reinforcement learning is a fundamentally different approach from the supervised learning and unsupervised learning mentioned above. Reinforcement learning means learning to maximize the reward obtained when taking an action in a given situation. Supervised learning and unsupervised learning analyze patterns or characteristics of given data to make decisions about new data, but reinforcement learning aims only to maximize rewards without identifying patterns or characteristics of given data. At this time, the learner (agent) does not receive any guidance on what action to take, but simply interacts with the given environment and aims to find out how to maximize the reward through trial and error.

The important point here is that when performing a specific action, one should not only consider the direct reward for that action, but also consider the reward that has a continuous effect in the subsequent situation after performing the specific action. AlphaGo, developed by Deepmind, is a representative example of applying a representative reinforcement learning algorithm [9]. If we assume that Baduk is learned through supervised learning, we must define what action to take for every situation, but the number of cases that exist in Baduk is almost infinite, making it impossible to apply supervised learning. However, reinforcement learning learns itself to maximize rewards by adjusting the probability of the policy network based on numerous trials and errors through past Baduk records.

There have been various studies attempting to apply machine learning to scheduling. Existing studies on schedule planning using machine learning that have been conducted so far are as follows. Jae-seok Heo developed a decision-making model to increase operation rate in the semiconductor packaging process using a deep neural network (DNN) [10], and Cho Yong-cheol used an artificial neural network (Artificial Neural Network) to improve wafer processing in the semiconductor manufacturing process. An optimal scheduling method for movement paths was proposed [11], and Jinyoung Kim proposed a DQN packet scheduling algorithm in wireless networks using the DQN (Deep-Q-Network) algorithm [12].

Ae-Kyung Kim conducted a study [13] to select an appropriate heuristic principle considering the current scheduling state using reinforcement learning. A decision-making model that determines the task sequence on a single machine was developed using Q-learning [14], and Hyun-Seok Jeong proposed a job shop scheduling method with various lengths of tasks and various requirements using reinforcement learning [14]. 15], Hyunjun Shin proposed an algorithm that can dynamically change job input policies over time based on reinforcement learning [16]. Thomas viewed the job shop scheduling problem as a sequential decision-making problem and developed a reinforcement learning model that independently determines allocation [17], and Jamal implemented a scheduling method that considers random task arrival and machine fixation using reinforcement learning [18].

Among the various existing studies on machine learning described above, it has recently been widely used and studied in the field of production management using reinforcement learning techniques, especially in the field of scheduling. The following describes representative studies that apply reinforcement learning to the field of scheduling. The Markov process has been widely adopted to solve control problems in situations where there are sudden changes in system dynamics [19]. For example, the Markov decision process (MDP), a fuzzy system, is advantageous for solving decision-making problems where the system dynamics are unknown [20].

MDP consists of state, action, reward at the current point, next state, and transition probability function to the next state. In MDP, as an action is executed, a transition is made from the current state to the next state, and the reward for the result is observed. At this time, the time interval from the current state to the next state is defined as the period. With the advantage of being able to learn MDP policies without explicitly knowing the model, reinforcement learning can be used for various decision-making problems modeled with MDP. Reinforcement learning aims to learn through interactions between the agent and its external environment, which encompasses everything. Among them, Q-learning is widely adopted as a representative model-free reinforcement learning approach [21].

After the manufacturing system was formalized as MDP to solve the scheduling problem, reinforcement learning was introduced to learn the policy. Given observed states from the external environment corresponding to the manufacturing system, the agent makes scheduling decisions by predicting the estimated value (Q) of each action. In manufacturing line scheduling, an action refers to a scheduling decision defined by a heuristic technique or a list of candidate jobs. Many existing studies have adopted a method of defining priority index-based heuristic techniques as actions [22]. By applying this definition method, a behavioral definition that mixes several dispatching rules such as earliest due date (EDD) and shortest processing time (SPT) was used for jobshop scheduling problems [23]. By maintaining the learned reinforcement learning-based rules for a specific period of time, we were able to successfully respond to dynamic environments. More recently, the variability of delivery times and production requirements has also been considered [24]. Another class of actions is defined as a list of candidate jobs suitable for the current situation [25]. Alternatively, it may be the top K EDD jobs, job attributes, or job product type [26, 22, 27, 28]. In some studies, action refers to the policy of changing schedules created by other techniques. For example, the agent learns how to select hyperparameters to be used in simulated annealing techniques [29], VNS [30], and cuckoo search [31].

In a series of studies targeting parallel facility scheduling problems [32, 33], the Q-learning technique was introduced for the purpose of minimizing critical delay time. In these studies, the Q value was approximated using a linear basis function, the states of all jobs and equipment were expressed in state features, six heuristics were designed as actions, and reinforcement learning techniques were used to dynamically select rules to create individual heuristics. was able to surpass The previous study [32] dealt with order-dependent setup, but the agent was evaluated only on the same scheduling problem as during training, whereas the follow-up study [33] did not consider setup, but was evaluated on several problems other than the learning task.

Yuan et al. [34, 35] focused on considering preparation time constraints and facility failures with the total delay time and delay miscellaneous minimization objective functions, respectively, and the Q value approximation involves storing the Q values of the explored state-action pairs in a table. technique was used. In this technique, it was necessary to separate attributes with continuous values into several groups in order to express them in a limited-sized table. For example, the average lateness value of a job was divided into two types, large or small, based on 0. The status was expressed separately [35].

Due to its complexity and usefulness in industry, the job shop scheduling problem has been studied using reinforcement learning for a long time [36]. Bouazza et al. [68] used a table-based Q-learning technique in a flexible job-shop scheduling problem under delivery constraints. A study was conducted to apply reinforcement learning by considering resource constraints in the semiconductor test process, which is a job-shop scheduling problem with setup [38]. Meanwhile, there is also research that addresses both scheduling and queuing problems by combining queuing theory with reinforcement learning-based scheduling research [39].

III. LIMITATIONS OF EXISTING STUDIES

As explained earlier, there are many studies using treadmills, especially reinforcement learning theory, in the field of production management. However, the following limitations exist in existing studies. First, there is data dependency. Many machine learning and reinforcement learning models rely on large amounts of high-quality data. In production management and scheduling, collecting and cleaning this data can be difficult or expensive. This can limit the efficiency and accuracy of the model.

Second, the complexity of the model and the difficulty of interpretability. Complex models can provide high prediction accuracy, but their internal mechanisms are difficult to understand and interpret. Production processes and scheduling often involve decisions from multiple stakeholders, so being able to clearly understand and explain the decisions in your model is important.

Third, there are difficulties with real-time optimization. Production environments and schedules are constantly changing, and optimal decisions often need to be made in real time. However, many machine learning and reinforcement learning models have limitations in supporting real-time data processing and immediate decision-making.

Lastly, there is the difficulty of general application. It is difficult to apply a model optimized for a specific factory or production line to other environments. Differences in production processes, product diversity, and changes in production goals make generalization of the model difficult.

IV. CONCLUSION

The application of machine learning and reinforcement learning in the field of production management and scheduling is very promising, but continuous efforts are needed to overcome the limitations mentioned above and improve research directions. These efforts will allow us to develop more efficient and reliable production systems and scheduling methodologies.

First, the development of data-efficient learning methods is necessary. We need to focus on developing models that can learn effectively even with small amounts of data. Methods such as transfer learning and few-shot learning may be promising.

Second, interpretable AI technology must be applied. We need to expand the application of interpretable AI technologies that can understand and explain the decision processes of models. This can increase stakeholder trust and increase acceptance of the model's decisions.

Third, the development of real-time optimization technology is essential. Advances in reinforcement learning and machine learning technologies that support real-time data processing and decision-making are needed. For this purpose, technologies such as streaming data processing and online learning are important.

Lastly, we will need to proceed with the development of generalized and customized solutions. In addition to developing models that can be applied to a variety of production environments and processes, attention should also be paid to developing solutions that can be customized to specific conditions. For this purpose, meta learning or modular design approaches can be useful.

This research received support from the Ministry of SMEs and Startups of the Republic of Korea in 2024.

V. REFERENCES

[1] Z. Chen, R. Ding, T. W. Chin and D. Marculescu, "Understanding the Impact of Label Granularity on CNN-Based Image Classification." 2018 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 895-904, 2018.

[2] M. O. Parkhi, A. Vedaldi, A. Zisserman and C. V. Jawahar, "Cats and dogs." 2012 IEEE conference on computer vision and pattern recognition. IEEE, 2012.

[3] H. Gao, B. Cheng, J. Wang, K. Li, J. Zhao and D. Li, "Object Classification Using CNN Based Fusion of Vision and LIDAR in Autonomous Vehicle Environment." IEEE Transactions on Industrial Informatics, Vol. 14, No. 9, pp. 4224-4231, 2018.

[4] D. C. Ciresan, U. Meier, L. M. Gambardella and J. Schmidhuber, "Convolutional Neural Network Committees for Handwritten Character Classification." 2011 International Conference on Document Analysis and Recognition, pp. 1135-1139, 2011.

[5] R. Gunn and Steve, "Support vector machines for classification and regression." ISIS technical report 14.1, pp. 5-16, 1998.

[6] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-Means Clustering Algorithm." Journal of the Royal Statistical Society. Series C (Applied Statistics), Vol. 28, No. 1, pp. 100-108, 1979.

[7] D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial–temporal data."Data & Knowledge Engineering, Vol. 60, No. 1, pp. 208-221, 2007.

[8] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction." MIT press, 2018.

[9] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller, "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602, 2013.

[10] Jaeseok Huh, Inbeom Park, Seongmin Lim, Bohyung Paeng, Jonghun Park and Kwanho Kim, ,"Learning to Dispatch Operations with Intentional Delay for Re-Entrant Multiple-Chip Product Assembly Lines" Information Technology and Architecture", MDPI, Open Access Journal, Vol.10 No.11, pp.1-21, 2018.

[11] Yongchul Cho, Hyeonchan Cho, Jongwon Kim, Liang Zhang and Hongtae Jeon, "A Study on Optimal Scheduling of Multi-Spinner's Manufacturing Process Using Artificial Neural Network," International Journal of Fuzzy Logic and Intelligent Systems , Vol.18, No.1, pp.157-160, 2008.

[12] Jinyoung Kim, "Packet scheduling algorithm using deep Q-Network in wireless network", M.S Thesis, Yonsei University, 2018.

[13] Aekyung Kim, "Ensemble-based Quality Classification and Deep Reinforcement Learning-based Production Scheduling ", Ph.D. Thesis, KyungHee University, 2018.

[14] Jangha Kang, "Feasibility Study on the Reinforcement Learning for a Dynamic Scheduling Problem", Journal of advanced engineering and technology, Vol.10, No.4, pp. 453-456, 2017.

[15] Hyunseok Jeong, Minwoo Kim, Byungjun Lee, Kyoungtae Kim and Heeyong Youn, "An Neural Network Approach to Job-shop Scheduling based on Reinforcement Learning", Proceedings of the Korean Society of Computer Information Conference, Vol. 26, No. 2, pp.47-48, 2018.

[16] Hyunjoon Shin, Jae-Pil Ru and Jaewoo Lee, "Adaptive scheduling algorithm for manufacturing process with nonstationary rework probabilities using reinforcement learning", Journal of academia-industrial technology, Vol.29, Iss.3, pp.1180-1180, 2010.

[17] Gabel Thomas and Riedmiller Martin, "Distributed policy search reinforcement learning for job-shop scheduling tasks.", International Journal of Production Research, Vol.50, Iss.1, pp.41-61, 2012.

[18] Jamal Shahrabia, Mohammad Amin Adibib and Masoud Mahootchia, "A reinforcement learning approach to parameter estimation in dynamic job shop scheduling", Computers and Industrial Engineering, Vol.110, pp.75-82, 2017.

[19] O. L. Costa and M. D. Fragoso, \Discrete-time lq-optimal control problems for innite markov jump parameter systems," IEEE Transactions on Automatic Control, vol. 40, no. 12, pp. 2076-2088, 1995.

[20] M. L. Puterman, Markov decision processes: discrete stochastic dynamic programming. John Wiley & Sons, 2014.

[21] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction. MIT press, 2018.

[22] T. Gabel, and M. Riedmiller, Adaptive reactive job-shop scheduling with reinforcement learning agents, International Journal of Information Technology and Intelligent Computing, vol. 24, no. 4, pp. 14-18, 2008.

[23] Y.-R. Shiue, K.-C. Lee, and C.-T. Su, Real-time scheduling for a smart factory using a reinforcement learning approach, Computers & Industrial Engineering, vol. 125, pp. 604-614, 2018.

[24] Y.-R. Shiue, K.-C. Lee, and C.-T. Su, A reinforcement learning approach to dynamic scheduling in a product-mix exibility environment, IEEE Access, vol. 8, pp. 106542-106553, 2020.

[25] S. Qu, J. Wang, and G. Shivani, Learning adaptive dispatching rules for a manufacturing process system by using reinforcement learning approach, in 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, pp. 18, 2016.

[26] B. Waschneck, A. Reichstaller, L. Belzner, T. Altenmuller, T. Bauernhansl, A. Knapp, and A. Kyek, Deep reinforcement learning for semiconductor production scheduling, in 2018 29th annual SEMI advanced semiconductor manufacturing conference (ASMC). IEEE, pp. 301-306, 2018.

[27] H. Rummukainen and J. K. Nurminen, Practical reinforcement learning experiences in lot scheduling application, IFAC-PapersOnLine, vol. 52, no. 13, pp. 1415-1420, 2019.

[28] I.-B. Park, J. Huh, J. Kim, and J. Park, A reinforcement learning approach to robust scheduling of semiconductor manufacturing facilities, IEEE Transactions on Automation Science and Engineering, vol. 17, no. 3, pp. 1420-1431, 2019.

[29] W. Zhang, and T. G. Dietterich, A reinforcement learning approach to job shop scheduling, in IJCAI, vol. 95. Citeseer, pp. 1114-1120, 1995.

[30] J. Shahrabi, M. A. Adibi, and M. Mahootchi, A reinforcement learning approach to parameter estimation in dynamic job shop scheduling, Computers & Industrial Engineering, vol. 110, pp. 75-82, 2017.

[31] Z. Cao, C. Lin, M. Zhou, and R. Huang, Scheduling semiconductor testing facility by using cuckoo search algorithm with reinforcement learning and surrogate modeling, IEEE Transactions on Automation Science and Engineering, vol. 16, no. 2, pp. 825-837, 2018.

[32] Z. Zhang, L. Zheng, and M. X. Weng, Dynamic parallel machine scheduling with mean weighted tardiness objective by q-learning, The International Journal of Advanced Manufacturing Technology, vol. 34, no. 9-10, pp. 968-980, 2007.

[33] Z. Zhang, L. Zheng, N. Li, W. Wang, S. Zhong, and K. Hu, Minimizing mean weighted tardiness in unrelated parallel machine scheduling with reinforcement learning, Computers & operations research, vol. 39, no. 7, pp. 1315-1324, 2012.

[34] B. Yuan, L. Wang, and Z. Jiang, Dynamic parallel machine scheduling us ing the learning agent, in 2013 IEEE international conference on industrial engineering and engineering management. IEEE, 2013, pp. 1565-1569.

[35] B. Yuan, Z. Jiang, and L. Wang, Dynamic parallel machine scheduling with random breakdowns using the learning agent, International Journal of Services Operations and Informatics, vol. 8, no. 2, pp. 94-103, 2016.

[36] J. Hong, and V. V. Prabhu, Distributed reinforcement learning control for batch sequencing and sizing in just-in-time manufacturing systems, Applied Intelligence, vol. 20, no. 1, pp. 71-87, 2004.

[37] W. Bouazza, Y. Sallez, and B. Beldjilali, A distributed approach solving partially flexible job-shop scheduling problem with a q-learning effect, IFAC Papers OnLine, vol. 50, no. 1, pp. 15890-15895, 2017.

[38] Z. Zhang, L. Zheng, F. Hou, and N. Li, Semiconductor final test scheduling with sarsa algorithm, European Journal of Operational Research, vol. 215, no. 2, pp. 446-458, 2011.

[39] D. Terekhov, T. T. Tran, D. G. Down, and J. C. Beck, Integrating queueing theory and scheduling for dynamic scheduling problems, Journal of Artificial Intelligence Research, vol. 50, pp. 535-572, 2014.